Opening Up Science

Marcus R. Munafò [1,2]

1. MRC Integrative Epidemiology Unit (IEU) at the University of Bristol, Bristol, United Kingdom.
2. UK Centre for Tobacco and Alcohol Studies, School of Experimental Psychology, University of Bristol, Bristol, United Kingdom.

Corresponding author: Marcus R. Munafò, School of Experimental Psychology, University of Bristol, Bristol BS8 1TU, United Kingdom. T: +44.117.9546841; F: +44.1179288588; E: marcus.munafo@bristol.ac.uk

Abstract

Open Science is an umbrella term that encompasses making study materials, data, results and publications freely available. This not only enables wider dissemination of research findings (including to non-academics), but also promotes greater transparency and may improve the robustness and reproducibility of published research.

Keywords: Open Science; Open Access; Open Data; Pre-Registration; Reproducibility.

Opening Up Science

There is growing consideration of the possibility that many published research findings may be false (1). While cases of fraud and data fabrication do exist, arguably a greater problem is the wider incentive structures in science that reward discovery and novelty over replication. This in turn may encourage various behaviours that reduce the likelihood that a research finding will prove robust, such as running a series of small, underpowered, but ultimately publishable studies, or conducting a large number of statistical analyses and only reporting those that produce the most interesting results. For example, the average statistical power of individual studies has been estimated to be as low as 20% in some domains (2), while the introduction of the requirement that clinical trial primary outcomes be pre-registered prior to publication was associated with a reduction in the number of trials that showed a significant benefit of intervention on the (reported) primary outcome (3). While there is ongoing debate as to the nature and extent of the reproducibility "crisis" (as it has become known), it is timely to reflect on whether the process of scientific research can be improved.

A number of potential ways to improve reproducibility have been proposed (such as pre-registration of studies and analysis plans) (4, 5), many of which form part of what has become known as the Open Science movement. This encourages scientists to make their materials, data, and publications freely available to all. In its broadest sense, it includes open source software, open peer review (such as that practiced by the *Frontiers* family of journals) (6), and other resources (such as educational materials). A number of third-party services now exist to support Open Science, such as the Open Science Framework (OSF), a free, open-source service provided by the non-profit Center for Open Science (http://centerforopenscience.org). However, the principal motivation of the Open Science movement is not reproducibility but rather to promote wider access to the products of scientific

research (much of which is ultimately funded by public money or charitable donations), greater efficiency (through the sharing of materials and data), and improved quality control (through the ability to re-analyze data independently, but also as a result of the natural tendency to check one's own data one more time if these are to be made public!).

In addition to the advantages of adopting an Open Science model to science (e.g., increased efficiency through the sharing of materials and data) and the wider community (e.g., free access to research outputs), there are potential benefits to individual scientists and research groups. For example, it may encourage the harmonization of procedures within (and between) research groups, and improve quality control procedures (7). Nevertheless, adopting an Open Science approach can be a substantial undertaking, and may entail changes in procedures, the use of a number of platforms to make materials, data, and publications publicly available, and discussions with institutional ethics committees and research governance teams. Critically, certain aspects of Open Science may not be appropriate in some settings. For example, if there is a risk of participant identification in an anonymised data set (e.g., where the sample is drawn from a small, distinctive clinical population), particularly where sensitive information is involved, making data publicly available will not be appropriate. Pre-registration of study protocols may not be appropriate for exploratory research, or for secondary analyses of existing data. Any general move in the scientific community towards an Open Science model will necessarily be gradual, and will need to with genuine concerns that may exist about certain aspects of the model in specific settings.

Recently the OSF introduced Transparency and Openness Promotion (TOP) Guidelines for journals and publishers (http://cos/io/top). These introduce eight standards that encourage greater openness - in brief, they cover citation standards, data, analytic methods (code), research materials, design and analysis, pre-registration of studies, pre-registration of analysis plans, and replication (8). Not all of

these standards are applicable to all journals or disciplines, and therefore three levels for each standard are defined, with Level 1 presenting the fewest barriers to adoption (e.g., simply stating in the text of an article whether data are available, and if so where) and Level 3 representing the most stringent standards (e.g., mandating the deposition of data to a trusted repository and reported analyses reproduced independently prior to publication). Critically, these levels also allow for various aspects of the Open Science model to be adopted gradually, and allow journals to only adopt the guidelines up to a level that is appropriate in that field.

The principles of Open Science movement are not new – calls to improve the accessibility of data go back many years (9), while pre-registration is now the norm for clinical trials. However, if Open Science is to become the norm this will require a cultural change that will come about through both top-down and bottom-up activity. The former includes funders and publishers – for example, research funders are increasingly mandating data sharing and open access publication, while a number of journals (including *Addiction*) are signatories to the TOP Guidelines (8). At the same time, individual researchers and research groups can work to promote this cultural change, through training of early career researchers and the promotion of the principles of Open Science (7). It will also require a wider discussion within the scientific community of the potential advantages of the Open Science model, as well as barriers to adoption, including cases where exceptions will need to be made, and what new approaches and platforms are required to support it.

Acknowledgements

MRM is a member of the UK Centre for Tobacco and Alcohol Studies, a UKCRC Public Health Research: Centre of Excellence. Funding from British Heart Foundation, Cancer Research UK, Economic and Social Research Council, Medical Research Council, and the National Institute for Health Research, under the auspices of the UK Clinical Research Collaboration, is gratefully acknowledged.

References

1. Ioannidis JP. Why most published research findings are false. PLoS medicine. 2005 Aug;2(8):e124. PubMed PMID: 16060722. Pubmed Central PMCID: 1182327.

2. Button KS, Ioannidis JP, Mokrysz C, Nosek BA, Flint J, Robinson ES, et al. Power failure: why small sample size undermines the reliability of neuroscience. Nature reviews Neuroscience. 2013 May;14(5):365-76. PubMed PMID: 23571845.

3. Kaplan RM, Irvin VL. Likelihood of Null Effects of Large NHLBI Clinical Trials Has Increased over Time. PloS one. 2015;10(8):e0132382. PubMed PMID: 26244868. Pubmed Central PMCID: 4526697.

4. Munafo M, Noble S, Browne WJ, Brunner D, Button K, Ferreira J, et al. Scientific rigor and the art of motorcycle maintenance. Nature biotechnology. 2014 Sep;32(9):871-3. PubMed PMID: 25203032.

5. Ware JJ, Munafo MR. Significance chasing in research practice: causes, consequences and possible solutions. Addiction. 2015 Jan;110(1):4-8. PubMed PMID: 25040652.

6. Poschl U. Multi-stage open peer review: scientific evaluation integrating the strengths of traditional peer review with the virtues of transparency and self-regulation. Frontiers in computational neuroscience. 2012;6:33. PubMed PMID: 22783183. Pubmed Central PMCID: 3389610.

7. Attwood AS, Munafo MR. Navigating an open road. Journal of clinical epidemiology. 2015 Jul 7. PubMed PMID: 26163125.

8. Nosek BA, Alter G, Banks GC, Borsboom D, Bowman SD, Breckler SJ, et al. Promoting an open research culture. Science. 2015 Jun 26;348(6242):1422-5. PubMed PMID: 26113702.

9. Davey Smith G. Increasing the accessibility of data. Bmj. 1994 Jun 11;308(6943):1519-20. PubMed PMID: 8019302. Pubmed Central PMCID: 2540522.